Assessing time dependent changes in microbial composition of biological crime scene traces using microbial RNA markers

Salzmann, Andrea Patrizia ; Arora, Natasha ; Russo, Giancarlo ; Kreutzer, Susanne ; Snipen, Lars ; Haas, Cordula

Abstract: Current body fluid identification methods do not reveal any information about the time since deposition (TsD) of biological traces, even though determining the age of traces could be crucial for the investigative process. To determine the utility of microbial RNA markers for TsD estimation, we examined RNA sequencing data from five forensically relevant body fluids (blood, menstrual blood, saliva, semen, and vaginal secretion) over seven time points, ranging from fresh to 1.5 years. One set of samples was stored indoors while another was exposed to outdoor conditions. In outdoor samples, we observed a consistent compositional shift, occurring after 4 weeks: this shift was characterized by an overall increase in non-human eukaryotic RNA and an overall decrease in prokaryotic RNA. In depth analyses showed a high fraction of tree, grass and fungal signatures, which are characteristic for the environment the samples were exposed to. When examining the prokaryotic fraction in more detail, three bacterial phyla were found to exhibit the largest changes in abundance, namely Actinobacteria, Proteobacteria and Firmicutes. More detailed analyses at the order level were done using a Lasso regression analysis to find a predictive subset of bacterial taxa. We found 26 bacterial orders to be indicative of sample age. Indoor samples did not reveal such a clear compositional change at the domain level: eukaryotic and prokaryotic abundance remained relatively stable across the assessed time period. Nonetheless, a Lasso regression analysis identified 32 bacterial orders exhibiting clear changes over time, enabling the prediction of TsD. For both indoor and outdoor samples, a larger number (around 60%) of the bacterial orders identified as indicative of TsD are part of the Actinobacteria, Proteobacteria and Firmicutes. In summary, we found that the observed changes across time are not primarily due to changes associated with body fluid specific bacteria but mostly due to accumulation of bacteria from the environment. Orders of these environmental bacteria could be evaluated for TsD prediction, considering the location and environment of the crime scene. However, further studies are needed to verify these findings, determine the applicability across samples, replicates, donors, and other variables, and also to further assess the effect of different seasons and locations on the samples.

DOI: https://doi.org/10.1016/j.fsigen.2021.102537

Originally published at:
Salzmann, Andrea Patrizia; Arora, Natasha; Russo, Giancarlo; Kreutzer, Susanne; Snipen, Lars; Haas, Cordula (2021). Assessing time dependent changes in microbial composition of biological crime scene traces using microbial RNA markers. Forensic Science International. Genetics, 53:102537.
DOI: https://doi.org/10.1016/j.fsigen.2021.102537

Forensic Science International: Genetics 53 (2021) 102537

Contents lists available at ScienceDirect

Forensic Science International: Genetics

journal homepage: www.elsevier.com/locate/fsigen

Research paper

Assessing time dependent changes in microbial composition of biological crime scene traces using microbial RNA markers

Andrea Patrizia Salzmann [a], Natasha Arora [a], Giancarlo Russo [c], Susanne Kreutzer [c], Lars Snipen [b,1], Cordula Haas [a,*,1]

[a] Zurich Institute of Forensic Medicine, University of Zurich, Switzerland
[b] Department of Chemistry, Biotechnology and Food Science, Norwegian University of Life Sciences, Ås, Norway
[c] Functional Genomics Centre Zurich (FGCZ), University of Zurich/ETH Zurich, Switzerland

ARTICLE INFO

Keywords:
Forensics
Transcriptome analysis
Body fluids
Microbial composition
Time since deposition (TsD)

ABSTRACT

Current body fluid identification methods do not reveal any information about the time since deposition (TsD) of biological traces, even though determining the age of traces could be crucial for the investigative process. To determine the utility of microbial RNA markers for TsD estimation, we examined RNA sequencing data from five forensically relevant body fluids (blood, menstrual blood, saliva, semen, and vaginal secretion) over seven time points, ranging from fresh to 1.5 years. One set of samples was stored indoors while another was exposed to outdoor conditions. In outdoor samples, we observed a consistent compositional shift, occurring after 4 weeks: this shift was characterized by an overall increase in non-human eukaryotic RNA and an overall decrease in prokaryotic RNA. In depth analyses showed a high fraction of tree, grass and fungal signatures, which are characteristic for the environment the samples were exposed to. When examining the prokaryotic fraction in more detail, three bacterial phyla were found to exhibit the largest changes in abundance, namely Actinobacteria, Proteobacteria and Firmicutes. More detailed analyses at the order level were done using a Lasso regression analysis to find a predictive subset of bacterial taxa. We found 26 bacterial orders to be indicative of sample age. Indoor samples did not reveal such a clear compositional change at the domain level: eukaryotic and prokaryotic abundance remained relatively stable across the assessed time period. Nonetheless, a Lasso regression analysis identified 32 bacterial orders exhibiting clear changes over time, enabling the prediction of TsD. For both indoor and outdoor samples, a larger number (around 60%) of the bacterial orders identified as indicative of TsD are part of the Actinobacteria, Proteobacteria and Firmicutes. In summary, we found that the observed changes across time are not primarily due to changes associated with body fluid specific bacteria but mostly due to accumulation of bacteria from the environment. Orders of these environmental bacteria could be evaluated for TsD prediction, considering the location and environment of the crime scene. However, further studies are needed to verify these findings, determine the applicability across samples, replicates, donors, and other variables, and also to further assess the effect of different seasons and locations on the samples.

1. Introduction

Analysis of biological evidence typically begins with a preliminary screening for the presence of biological material. Body fluids are commonly identified by chemical analyses, immunological assays, protein catalytic activity tests, spectroscopic methods and microscopy [1,2]. New methods include mRNA profiling [3–8], tissue specific DNA methylation [9–11], proteomics [12] and microbial analysis [13–16].

Current body fluid identification methods do not reveal any information about the time since deposition (TsD) of biological traces. From a criminalistic point of view, estimation of when the crime was committed would be useful to determine the relevance of trace samples found at the scene, enable the verification of witnesses' statements, limit the number of suspects, and help corroborate alibis [17]. In recent years, several techniques for TsD determination of blood stains have been explored [17,18]. These techniques include spectroscopy, chromatography, and

* Corresponding author.
E-mail address: cordula.haas@irm.uzh.ch (C. Haas).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.fsigen.2021.102537
Received 22 October 2020; Received in revised form 21 May 2021; Accepted 22 May 2021
Available online 25 May 2021

electron spin resonance. These methods are promising for blood traces, but limited to colored stains and thus not easily transferred to other, white or nearly colorless traces such as saliva, semen, or vaginal fluid.

Several attempts have been made to estimate TsD based on RNA analyses. Most often, degradation patterns of different human messenger RNA (mRNA) and ribosomal RNA (rRNA) transcripts in blood were monitored [18–21]. Bauer et al. [19] showed that mRNA could be analyzed from blood samples which were 15 years old, and found that mRNA degradation levels differ significantly among samples of different ages. Anderson et al. [18,20] focused on transcripts of two housekeeping genes, ß-actin and 18S-rRNA, and found that the relative degradation patterns of these as examined through quantitative PCR (qPCR) could be used to estimate the age of blood stains. An optimized qPCR approach involving the quantification of small and long fragments within ß-actin mRNA and 18S-rRNA allowed to distinguish fresh from 6-day old blood samples, and 6-day from older blood samples. A subsequent study by Alshehhi et al. [22] focused on semen and saliva stains and found that mRNA molecules specific to these body fluids showed a unique degradation pattern over a period of one year, whereas miRNAs and the U6 reference gene were shown to be stable. The relative expression ratio was suggested as a potential method for TsD estimation.

Most notably, the introduction of massively parallel sequencing (MPS) has enabled the shift from utilizing small numbers of RNA markers to exploring and using the full range of RNA sequencing transcripts for TsD estimation. However, the low quantity and quality of RNA in forensic samples poses a genuine challenge in practice. Nevertheless, Lin et al. [23] successfully sequenced total RNA from forensically relevant body fluids (blood, menstrual blood, oral mucosa/saliva and vaginal secretion) and assessed the global gene expression levels over time. It was shown that sequencing reads with relatively good quality (85% of reads had a Q score >30) can be obtained from both fresh and aged (two and six weeks old) samples. Recently, Weinbrecht et al. [24] assessed the global abundance and degradation pattern of mRNA in blood, saliva, semen and vaginal secretion samples, across a one year interval. The abundance of the different transcripts decreased over time but at different rates.

In recent years, microbial forensics has emerged as a new research area [25–27]. Bacteria colonize vast areas of the human body such as skin, the gastrointestinal and urogenital tracts and the oral cavity. Overall, the number of human and bacterial cells in the human body are estimated to be similar [28,29]. However, human body sites display large variability in the density as well as the composition of microorganisms. In fact, studies analyzing the human microbiome have revealed that different body sites harbor very distinctive bacterial communities, allowing for body fluid identification based on microbial communities [26,30]. Several bacterial markers have been proposed for the identification of vaginal secretion [31–33], saliva [34,35] and feces [36].

Traditionally, hypervariable DNA regions of the 16S-rRNA gene are sequenced and used to analyze the bacterial composition of samples [13]. In order to distinguish body sites using such datasets, different statistical approaches have been proposed. For instance, Hanssen et al. [14] showed that a combination of principal component analysis (PCA) and linear discriminant analysis (LDA) can successfully be used to differentiate samples from saliva being deposited on skin and samples of skin only. In addition, Tackmann et al. [43] have shown that machine learning algorithms trained on large heterogeneous datasets provide high accuracy when predicting body fluids, not only from single source samples but also from mixtures generated *in-silico*. In another study, Hanssen et al. [15] developed a prediction model using partial least squares (PLS) in combination with LDA using data from the Human Microbiome Project for the identification of samples originating from the oral, nasal and vaginal cavity as well as skin and feces.

The human microbiome not only varies across body sites but also across individuals, across intrinsic conditions (e.g. diet, medication and diseases), and across environmental/extrinsic conditions [36–42, 44–46]. While intrinsic conditions have been explored extensively, little is known about the changes in microbial composition over time once a sample is exposed to the environment outside the human body. Since samples collected at crime scenes are subjected to the extracorporeal environment and are stored for some time before analysis, the composition of the microbial community might change with time, depending on environmental conditions. Dobay et al. [16] collected samples of forensically relevant body fluids (blood, saliva, skin, semen, menstrual blood and vaginal secretion) and assessed the microbial composition directly after collection and after one month. They found that characteristic signatures allowing the identification of the body site could be obtained even after prolonged storage time. Subsequent studies by Diez-Lopez et al. [47,48] also indicate the stability and reliability of such signatures over longer periods of time when the samples were stored at 4 °C as well as room temperature for up to seven years.

Microbial sequencing for body fluid prediction has already been applied to the investigation of two criminal cases in the Netherlands [49]. In these two cases, human STR profiling and microbial population profiling were applied to the same trace sample in order to simultaneously investigate the donor as well as the bodily origin of the sample. In both cases, the analyses were helpful in the court proceedings, and the suspect was eventually convicted.

In a recent study, we published an optimized RNA-Seq workflow for forensic samples (blood, menstrual blood, saliva, semen, skin and vaginal secretion) [50]. Body fluid specific bacterial signatures could be identified in fresh and aged samples. In addition, aged samples showed a higher level of RNA degradation and decreased bacterial diversity. The aim of this study was to gain further insights into the changes of microbial communities of forensically relevant body fluids over time and to determine the utility of microbial RNA markers for TsD estimation. Blood, menstrual blood, saliva, semen, and vaginal secretion samples were exposed to different environmental conditions (indoor, dry, room temperature versus outdoor, exposed to the environment but protected from rain). Total RNA was analyzed at seven different time points (time point 0/deposition, 1 day, 7 days, 4 weeks, 6 months, 1 year, 1.5 years). The microbial composition at different taxonomic levels was assessed across all time points, for indoor and outdoor conditions. To our knowledge, this is the first study evaluating microbial markers for TsD estimation and in particular microbial RNA markers. In a parallel manuscript, we assessed the degradation of human mRNA transcripts over time as an indicator of the time since deposition in biological crime scene traces [51].

2. Material and methods

2.1. Body fluid samples

Forensically relevant body fluids (blood, menstrual blood, saliva, semen, vaginal secretion) were collected on sterile cotton swabs (Milian, Nesselnbach, Switzerland). Menstrual blood and vaginal secretion were self-collected from the vagina directly on swabs by the donors. Semen and saliva were self-collected in sterile tubes, and 50 µl each were spotted on the swabs in the laboratory. Blood was collected by venipuncture into EDTA coated tubes and 50 µl were directly pipetted onto the swabs. The chemical EDTA is a chelating agent, shown to inhibit the growth of biofilms and disrupt gram-negative bacterial cell walls [52, 53]. It must be noted, however, that the samples were kept in the EDTA-coated tubes very briefly, before being pipetted onto swabs. Thus, we expect EDTA to have had a minimal influence on the microbial composition of the samples.

Three biological replicates were collected for each body fluid (i.e. from three individuals). In total 12 different donors participated in the study, each providing a sample from only one body site, with the exception of three participants. Of these three, two provided both saliva and menstrual blood samples, and one provided both venous blood and vaginal secretion. For blood, saliva and semen samples, entire swabs were utilized for the RNA extraction. For vaginal secretion and

menstrual blood samples, only ½ of the cotton tip of a swab was removed at a given sampling time point and used for RNA extraction. The samples were exposed to two different environmental conditions: (1) indoors, at room temperature in a dark and dry place (in a cupboard), (2) outdoors on the flat rooftop of the institute building (exposed to sun and wind but protected from rain). The samples were put in place in March–April for up to 1.5 years. A park with trees and a forest in close proximity surrounds the institute building. Residential areas with gardens are about 500 m away. RNA was extracted immediately after sample collection (time point 0/deposition) and after 1 day, 7 days, 4 weeks, 6 months, 1 year and 1.5 years. This sampling scheme resulted in a total of 210 samples: for each of the 5 body fluids, we had 3 different donors, 7 time points, and 2 environmental conditions. The sampling was approved by the local ethics commission (KEK), with a declaration of no objection (No. 24-2015).

Negative controls, comprising swabs without any sampled fluid, were also included: one sample processed right away without aging, and another two exposed for 4 weeks, in either indoor or outdoor conditions, resulting in a total of three negative controls.

2.2. RNA extraction and library preparation

During a previous study, we developed an RNA-Seq workflow, which was shown to be suitable to identify human as well as body fluid specific bacterial signatures [50]. The same workflow has been applied in the present study. The cotton tips of the swabs were peeled off the wooden handles using a new scalpel blade for each sample. RNA was extracted using the ReliaPrep™ RNA Cell Miniprep kit (Promega, Dübendorf, Switzerland). We used the protocol for $> 5 \times 10^5$ to 2×10^6 cells according to the manufacturer's instructions, with the following modifications: stains were incubated in lysis buffer for two to three hours at 56 °C. The final elution volume was 30 µl. RNA quantity was assessed using the QuantiFluor® RNA HS System (Promega) according to the manufacturer's protocol.

Total RNA libraries were constructed using the Trio RNA-Seq kit (Nugen, Leek, The Netherlands) according to the manufacturer's protocol. The kit is optimized for low input and challenging samples, suitable to generate sequencing libraries from total RNA (universal priming) [54]. No ribosomal depletion was performed. Fifty ng RNA in 10 µl was used as input volume; where this amount could not be reached, 10 µl extract was used regardless of concentration. Libraries were quality checked on the TapeStation 4200 (Agilent Technologies, Santa Clara, USA). Libraries generated from fresh, 1 day, 7 days and 4 week-old stains were sequenced on the Illumina HiSeq 4000 platform, while 6 months, 1 and 1.5 year old samples were sequenced on the Illumina NovaSeq 6000 platform. For the HiSeq 4000, the normalized sequencing libraries were clustered at 8 pM with the TruSeq SR Cluster kit on a v4-cBOT HS (Illumina, California, San Diego) and sequenced using the TruSeq SBS kit with 1×125 bp reads and a single barcode (8 bp). NovaSeq 6000 sequencing was performed with the NovaSeq 6000 reagent kits and flowcell (v1) with a starting concentration of 1.7 pM of the pooled sequencing libraries and a unique dual barcoding strategy generating 1×100 bp. Two different Illumina platforms were used because the Illumina HiSeq 4000 platform at our sequencing facility was replaced by the Illumina NovaSeq 6000 platform during the course of this project. The data quality of both sequencers are comparable. Single read sequencing is a standard method for RNA-Seq, especially for degraded samples where the size of sequencing libraries is relatively small (insert size ~150 bp).

2.3. Taxonomic profiles

Taxonomic assignment of the raw reads was conducted using kraken2 [55]. We used a customized database containing not only the genomes that are in the standard reference database (bacteria, archaea, viruses, human, Univec) but also the genomes of fungi, protozoa and plants (see the software documentation at https://github.com/DerrickWood/kraken2/wiki/Manual). This was done to map as many reads as possible, and to get an overview of the taxonomic composition in all samples. Reads mapping to the Univec library were discarded. The final read count table was normalized by dividing the read counts for each taxon per sample by the sum of read counts in the respective sample. Taxonomic bar plots were produced in R using Version 3.6.1.

2.4. Regression analysis

Further investigation of the prokaryotic RNA was conducted with regression analyses in order to identify potential bacterial taxa informative for TsD. The age of a sample was used as the response variable, and the taxonomic compositions at a chosen rank were used as predictors. First, taxa with average read counts below 100 were discarded since we have found that kraken2 often assigns a few reads to many taxa not present in the samples. Since the taxon signals are relative values, summing to 1.0 over all taxa, we transformed these data using the Centered Log-Ratio (CLR) transform, which is a standard procedure for such data [56]. We also added 100 pseudo counts to all read counts in this process, since the CLR-transform requires nonzero read counts. The value 100, both as the threshold for discarding very low abundance taxa and pseudo-counts, represents a very small value in this set of read counts, and should be modified accordingly for other data sets. The number of predictors, i.e. the number of taxa at the chosen taxonomic rank, may be large, and most of these are probably without any importance for the time effect. To search for a subset of relevant taxa, we employed the lasso method [57–60] for variable selection, implemented in the glmnet R-package. To avoid overfitting within the data, we conducted this analysis on all body-fluids together, i.e. looking for a generic time-effect, and carried out cross-validation, where samples of identical age and body-fluid were used as subsets.

3. Results

3.1. Data quality

All except two of the 210 sequenced samples reached the targeted 40–50 million reads per sample. These exceptions were a fresh semen sample and a 6-month-old vaginal secretion sample (indoor condition), with 6.98 million reads and 17.8 million reads, respectively. We speculate that the observed variation may be due to effects introduced during the laboratory workflow, not necessarily reflecting biological variation. Mapping reads with kraken2 and the customized database showed that 78–99.7% of reads in the indoor samples and 82–99.7% reads in the outdoor samples could be taxonomically assigned, with the largest number of unclassified reads in aged outdoor samples. Between 4% and 90% reads of the indoor as well as 0.09–81% reads of the outdoor samples were of human origin (Supplementary Figs. 1 and 2). Read counts for the negative controls were between one and six million reads per sample, much lower than that for body fluid samples (40–100 million reads) (data not shown).

3.2. Outdoor samples

We investigated the taxonomic composition of outdoor samples at the domain rank. Blood and semen samples displayed a dominance of eukaryotic reads throughout all time points, while this dominance became established in menstrual blood, vaginal secretion and saliva samples after the 4th week. This shift was accompanied by a general decrease in the prokaryotic fraction (Supplementary Figs. 3 and 4). When looking at time of deposition (time point 0), we found that up to 80% of total reads in blood and semen comprised human RNA. In the other body fluids, the percentage of human reads reached around 50% (menstrual blood, vaginal secretion) and 25% in saliva, at time of deposition. We thus excluded human reads in order to focus on the

effects of non-human eukaryotic RNA. Our analyses revealed a consistent and marked shift in all samples, with an increase in non-human eukaryotic RNA, that is, plants and fungi, as well as a decrease in prokaryotic RNA, occurring between 4 weeks and 6 months (Fig. 1). By 1 year, all samples displayed a dominance of environmental eukaryotic reads. In menstrual blood and vaginal secretion samples a donor effect in the taxonomic composition was detectable until 4 weeks of age (donor 3 in menstrual blood and donor 2 in vaginal secretion). The compositional shift to the dominance of eukaryotes was more gradual for semen and blood compared to the other samples.

We conducted further analyses on all eukaryotic reads, including human data, to identify the most frequent genera present (with total number of reads across samples above 1%) (Fig. 2). As observed earlier, in all body fluids the human fraction was dominant until 4 weeks. Subsequently, a shift from Homo to Quercus and Panicum was observed.

3.3. Prokaryotic composition of the outdoor samples

While the prokaryotic fraction exhibits a dramatic reduction over time, all samples nonetheless contained bacterial taxa. We expected this bacterial fraction to reflect exposure to the environment, with a potential increase in environmental bacterial RNA and decrease in endogenous bacterial RNA. We thus inspected the bacterial composition in greater detail, in order to identify potential patterns across time.

As observed in Fig. 3, the abundance of the phyla at the time of deposition varied across body fluids. Interestingly, however, all samples show an increase in the fraction of Actinobacteria across time (Fig. 3 and Supplementary Fig. 5). This increase is concurrent with a decrease in Proteobacteria in blood and semen samples. In contrast, in menstrual blood, vaginal secretion and saliva, the increase in Actinobacteria is associated with an increase in Proteobacteria and a decrease in Firmicutes.

3.4. Regression analysis using the prokaryotic fraction in the outdoor samples

To investigate which of the bacterial taxa were most indicative of TsD, we conducted a Lasso regression analysis at the order level to select a subset of the most influential bacteria (Fig. 4). A total of 26 taxa, that together showed high predictive value for TsD (regression coefficient of intercept (s0) = 279.65), were identified (Supplementary Table 1), of which around 65% were part of Actinobacteria, Proteobacteria and Firmicutes. Closer inspection of the taxa revealed that they are not associated with the body fluid but rather with the environment.

3.5. Indoor samples

Indoor samples displayed different patterns in the relative fractions of eukaryotic, prokaryotic and viral RNA, compared to outdoor samples. While blood and semen samples were dominated by eukaryotic RNA throughout the assessed time period, prokaryotes were predominantly observed in vaginal secretion, menstrual blood as well as saliva (Supplementary Fig. 4). The exceptions were one vaginal secretion sample and one menstrual blood sample, where eukaryotic RNA was dominant.

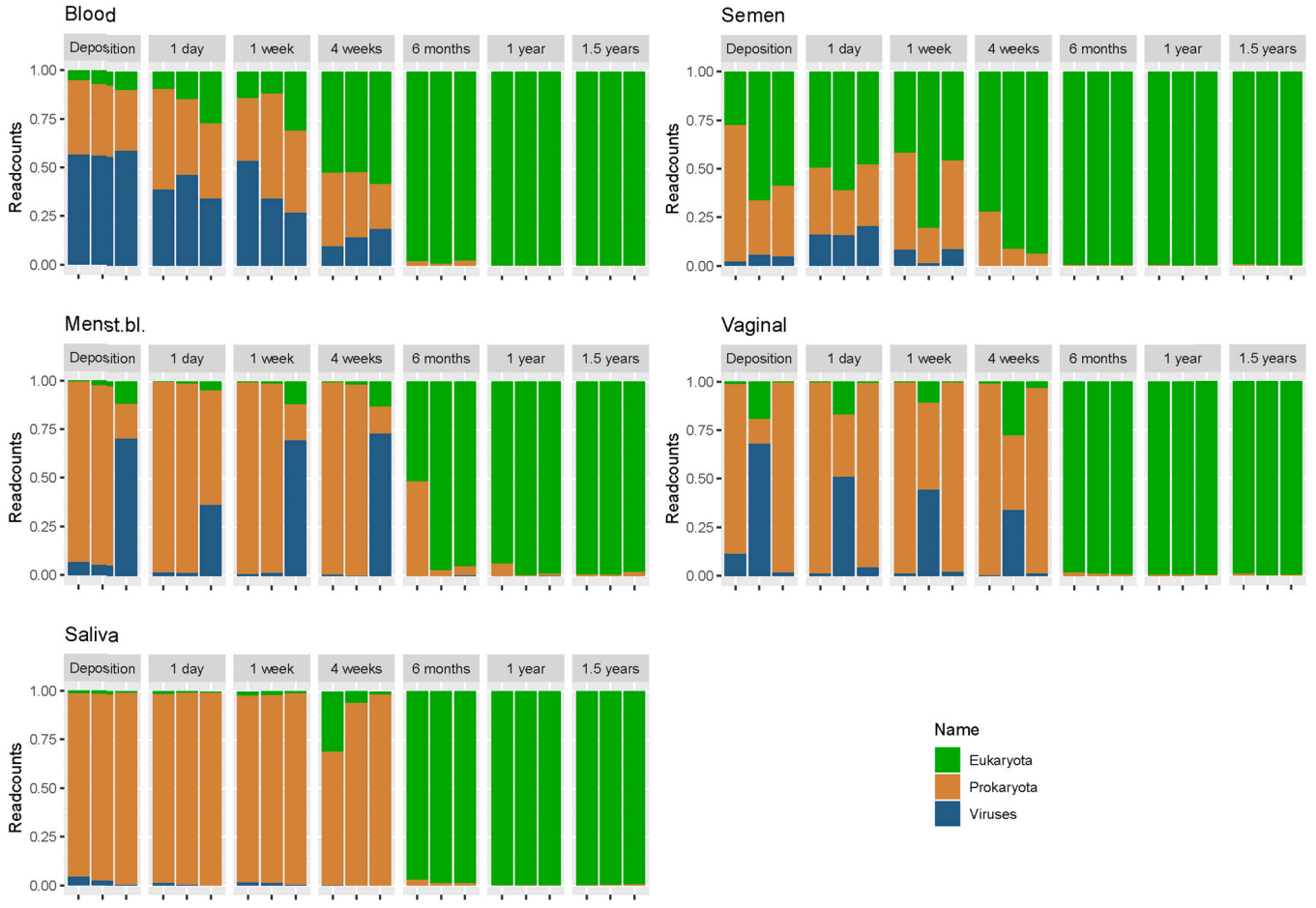

Fig. 1. Taxonomic bar plots of outdoor samples at domain level. The relative abundance of taxa is shown on the y-axis. On the x-axis, the samples for each of the three donors are shown, across the different time points. Human reads have been excluded. *(For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)*

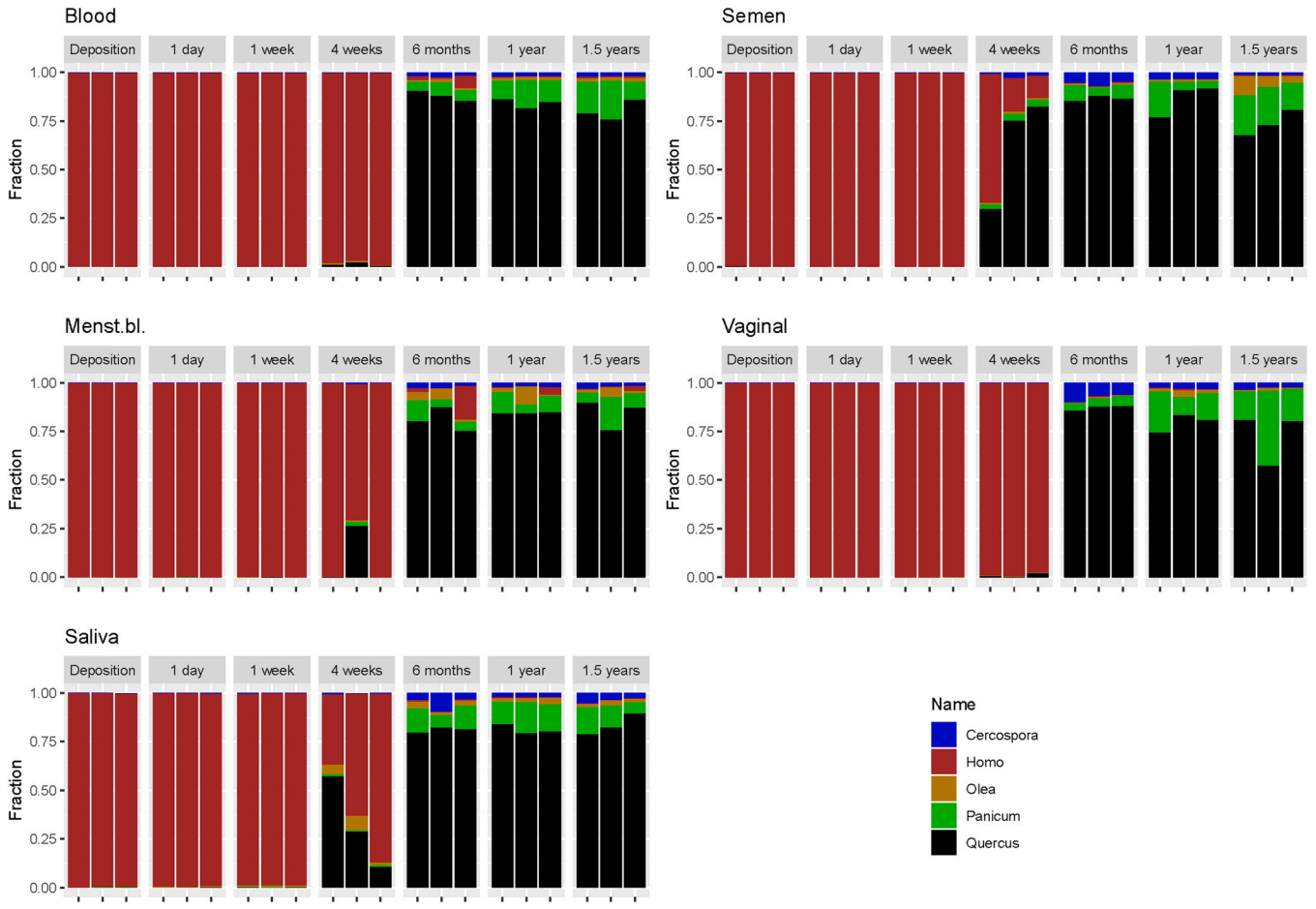

Fig. 2. Analysis of the eukaryotic genera composition of outdoor samples across time. Eukaryotic genera (with total number of reads across samples above 1%) are shown on the y-axis as a fraction of the total eukaryotic reads. On the x-axis, the samples for each of the three donors are shown, across the different time points. *(For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)*

As with the outdoor samples, we explored the non-human eukaryotic fraction further. After removal of human RNA reads, most samples exhibited less variation over time (Fig. 5). Particularly saliva samples consistently exhibited a very high number of prokaryotic reads throughout all time points. In blood and semen samples, the 6-month old samples stood out with a sudden increase in the prokaryotic fraction. This may represent a technical artifact rather than a biological effect. Data inspection revealed that rRNA counts for these samples were much higher than in samples previously prepared for the present study. The special amplification strategy in the Trio RNA-Seq kit (single primer isothermal amplification (SPIA) technology) reduces the rRNA content of the samples by preferentially priming and amplifying non-ribosomal RNA. While the kit had not expired after 6 months, it is possible that its prior usage and storage had an adverse effect on the reagents' quality and the subsequent sequencing outcome.

In the case of menstrual blood, for one donor, there was an increase in the eukaryotic composition at 1.5 years. A donor effect could be observed for donor 3 in menstrual blood and donor 2 in vaginal secretion. Otherwise, the samples displayed relatively stable compositions.

3.6. Prokaryotic composition of the indoor samples

The prokaryotic fraction of the indoor samples was assessed at the phylum level. We found that blood and semen were dominated by Proteobacteria (Fig. 6). After 1 year, Actinobacteria increased and started to dominate, while the abundance of Proteobacteria decreased. Menstrual blood and vaginal secretion were dominated by Firmicutes across all investigated time points. A donor effect could be observed for donor 1 and 3 in menstrual blood and donor 2 in vaginal secretion. Saliva showed a more diverse combination of phyla with only small changes across the 1.5 years.

3.7. Regression analysis using the prokaryotic fraction in the indoor samples

Compared to the outdoor samples, no distinct change in the domain-level taxa of indoor samples could be observed. We also used a Lasso regression analysis for the selection of a smaller subset of bacterial orders that are most indicative of TsD (Fig. 7, Supplementary Table 2). In total, 32 bacterial orders were found to be highly informative (regression coefficient of intercept (s0) = 149.87), of which 8 have also been identified in the outdoor samples. Approximately 62% of these bacterial orders are part of the Actinobacteria, Proteobacteria and Firmicutes phyla, which were also found to be indicative of TsD in outdoor samples. Closer inspection of the 32 orders revealed that these are mostly environmental rather than endogenous bacteria.

3.8. Negative controls

The negative controls yielded bacterial RNA reads, although a much smaller number compared to the body fluid samples. The bacteria identified in the indoor 4 week negative control are more characteristic of the laboratory environment than of any body fluid. Hence, we suspect that these bacteria were introduced on our samples by the laboratory

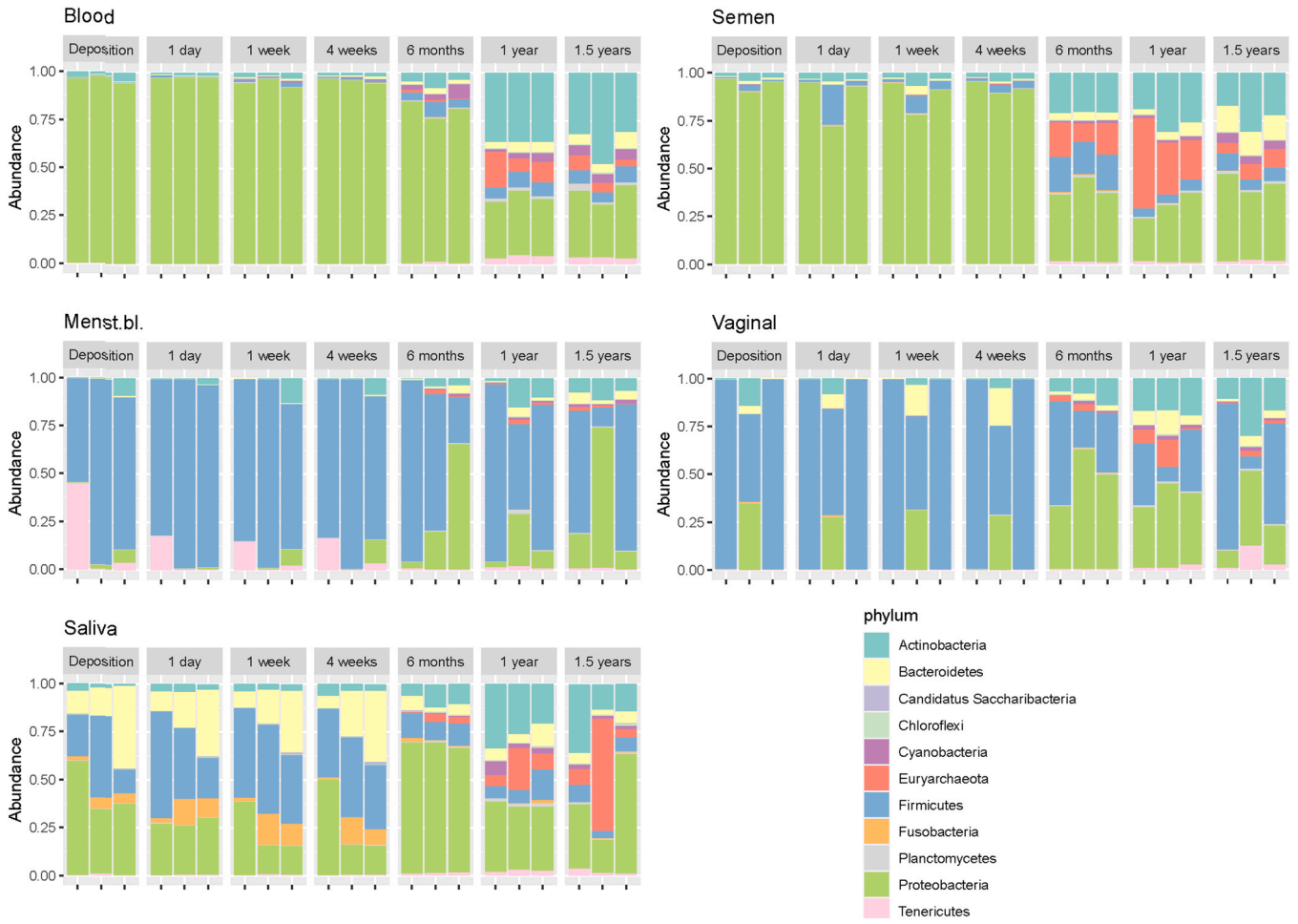

Fig. 3. Overview plots of outdoor samples at phylum level (prokaryotes only). The relative abundance of the different phyla is shown on the y-axis. On the x-axis, the samples for each of the three donors are shown, across the different time points. *(For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)*

environment. In the 4 week old outdoor negative control, we found signatures for Quercus (54% of eukaryotic reads), Panicum (22% of eukaryotic reads), Olea (13% of eukaryotic reads), Cercospora (3% of eukaryotic reads), confirming that these must have been introduced by air onto our samples.

3.9. STR analysis

It is worthwhile to note that TsD estimation is most useful for samples for which STR profiling is also possible. From the indoor samples retrieving full STR profiles is not expected to be problematic, even after 20 years or more [61,62]. However, the outdoor samples were subjected to harsh environmental conditions including storms, snow and UV light. Thus, we checked whether we could recover human DNA from a set of additional samples that were aged outdoors for even longer than the samples processed in this study (back-up samples), for a total of 2.5 years. We found that most samples (blood, menstrual blood, semen and vaginal samples) except for saliva yielded almost full STR profiles that matched the reference STR profiles of the different donors (data not shown).

4. Discussion

Determination of TsD is a desired addition to the current forensic toolbox and will aid investigators and courts of justice in assessing the relevance of biological traces. It is expected to contribute to a more accurate identification of true perpetrators as well as to clear the wrongfully accused. To date, there is no reliable method available to estimate the age of stains deposited at crime scenes despite efforts being made in this field of work. The aim of our study was to determine the utility of microbial RNA markers for TsD estimation. Hence, we examined RNA-Seq data from blood, menstrual blood, saliva, semen, and vaginal secretion over seven time points, ranging from fresh to 1.5 years. One set of samples was stored indoors, the other set was exposed to the outdoor environment. The microbial composition at different taxonomic levels was assessed across all time points for outdoor and indoor conditions.

Inspection of the taxonomic composition of both outdoor and indoor samples at the domain level (Figs. 1 and 5) as well as at the phylum level (Figs. 3 and 6) indicated that the five investigated body fluids can be categorized into three groups based on similarity: (a) blood and semen, (b) menstrual blood and vaginal secretion and (c) saliva. This finding is most likely due to the origin of the different body fluids and their respective microbial composition. For example, blood and semen from healthy individuals contain a minimal bacterial fraction, while the bacterial load of menstrual blood, vaginal secretion and saliva is high under normal conditions. Menstrual blood and vaginal secretion are more similar to each other due to their common body site of origin.

Overall, in outdoor samples there was a shift from prokaryotes to non-human eukaryotes between 4 weeks and 6 months (Fig. 1). This change is driven primarily by the increase of Quercus (Oak tree) and Panicum (switch grass) in all samples (Fig. 2). Oak trees are common in

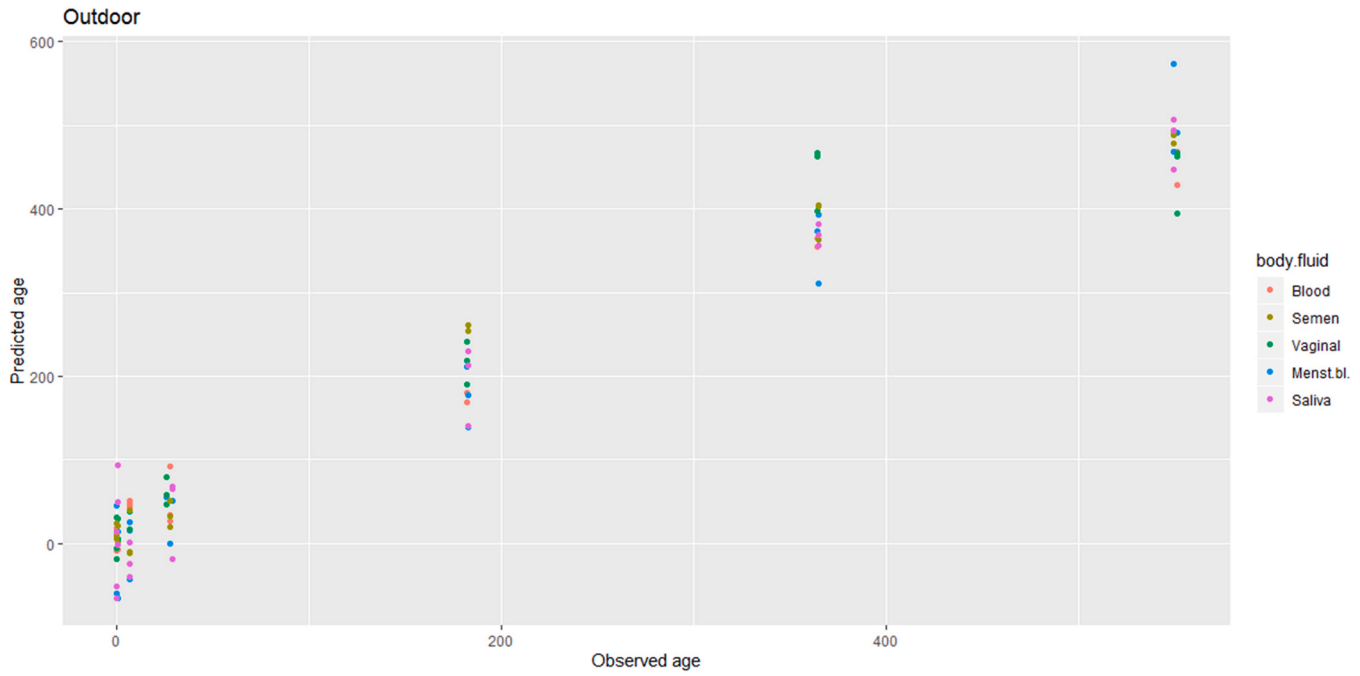

Fig. 4. Observed age of the outdoor samples (y-axis) versus predicted (fitted) age (x-axis). The plot is based on the 26 taxa at order rank selected by lasso to be good age predictors. *(For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)*

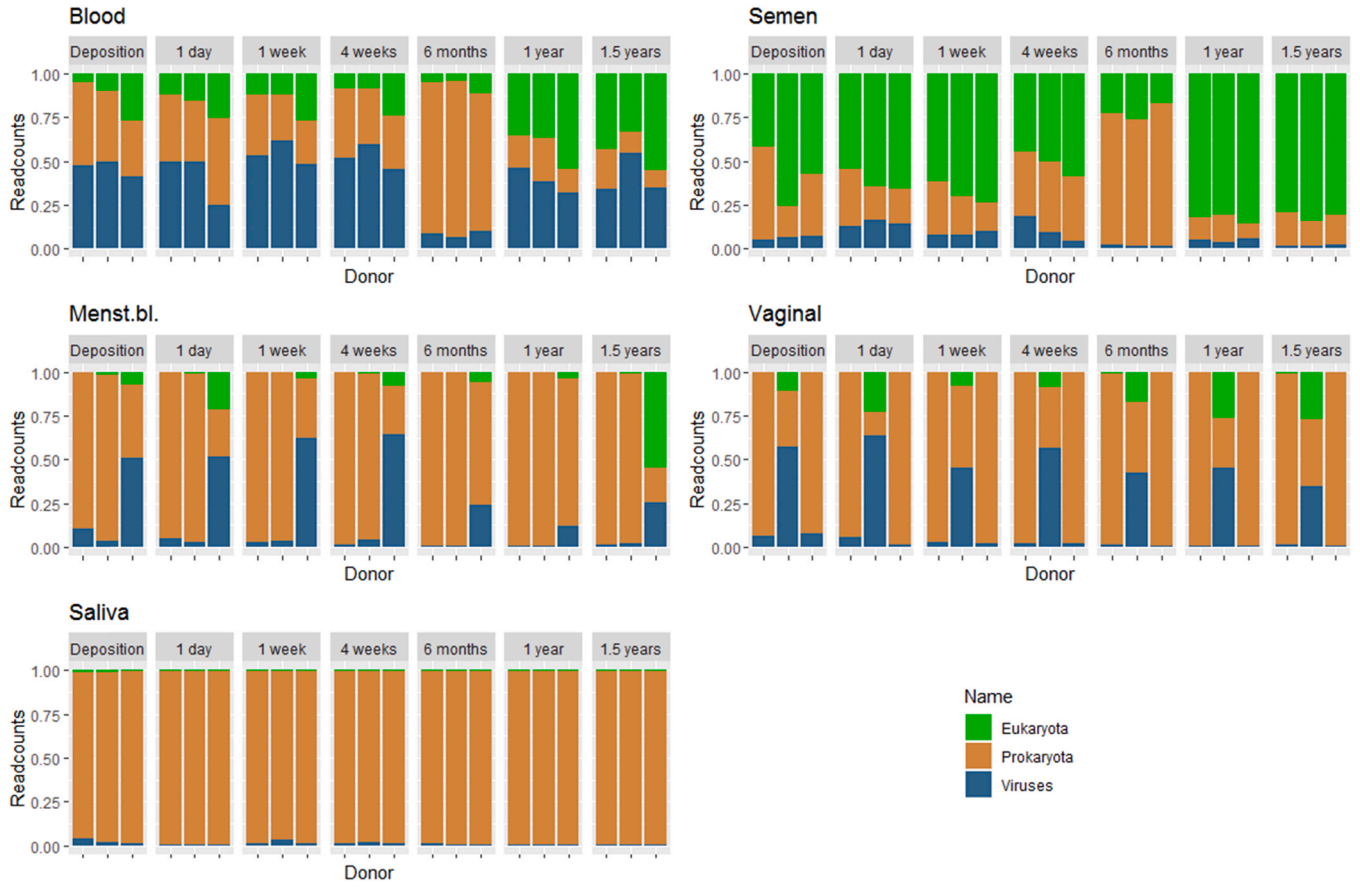

Fig. 5. Taxonomic bar plots of indoor samples at domain level. The relative abundance of taxa is shown on the y-axis. On the x-axis, the samples for each of the three donors are shown, across the different time points. Human reads have been excluded. *(For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)*

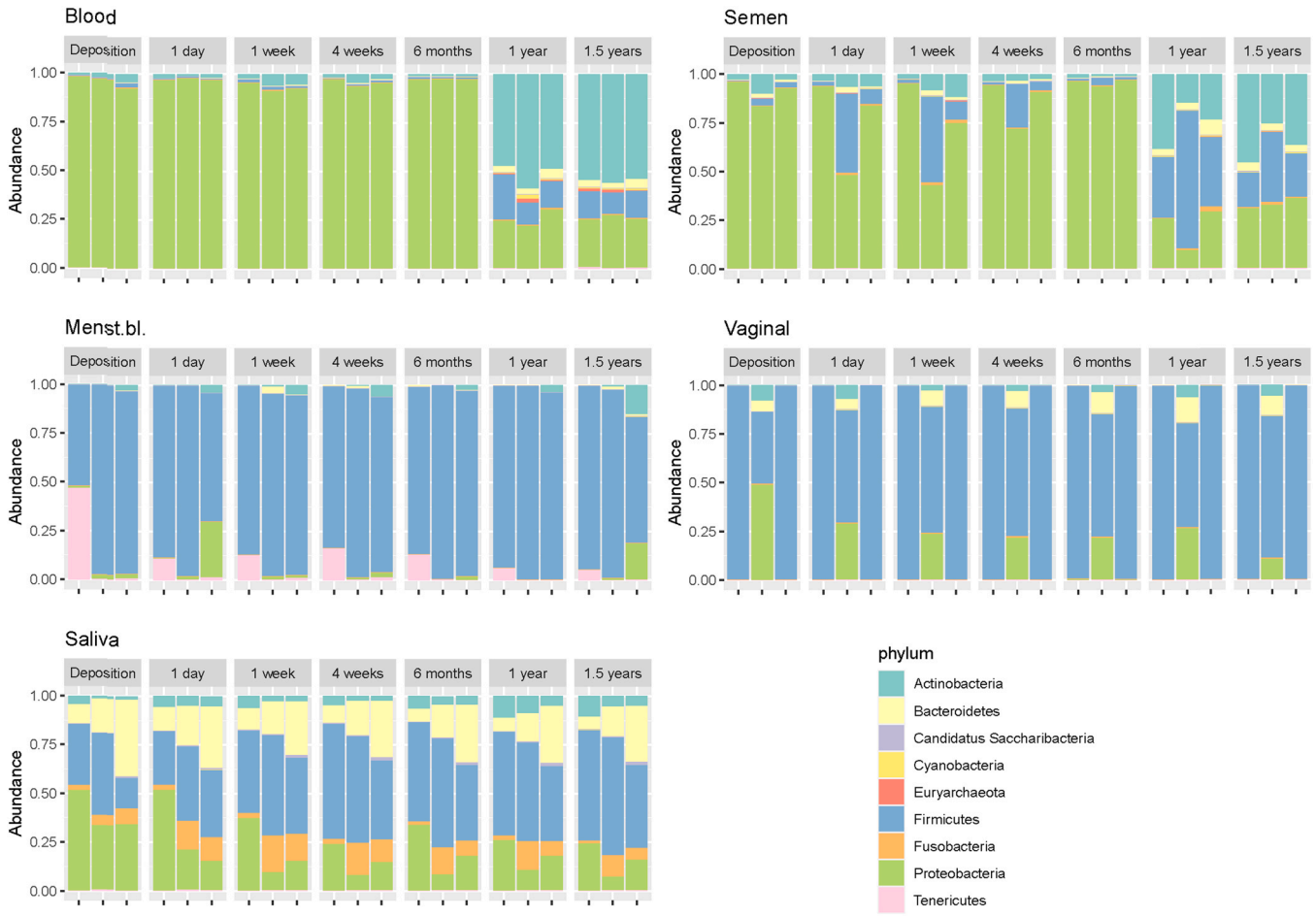

Fig. 6. Overview plots of indoor samples at phylum level (prokaryotes only). The relative abundance of the different phyla is shown on the y-axis. On the x-axis, the samples for each of the three donors are shown, across the different time points. *(For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)*

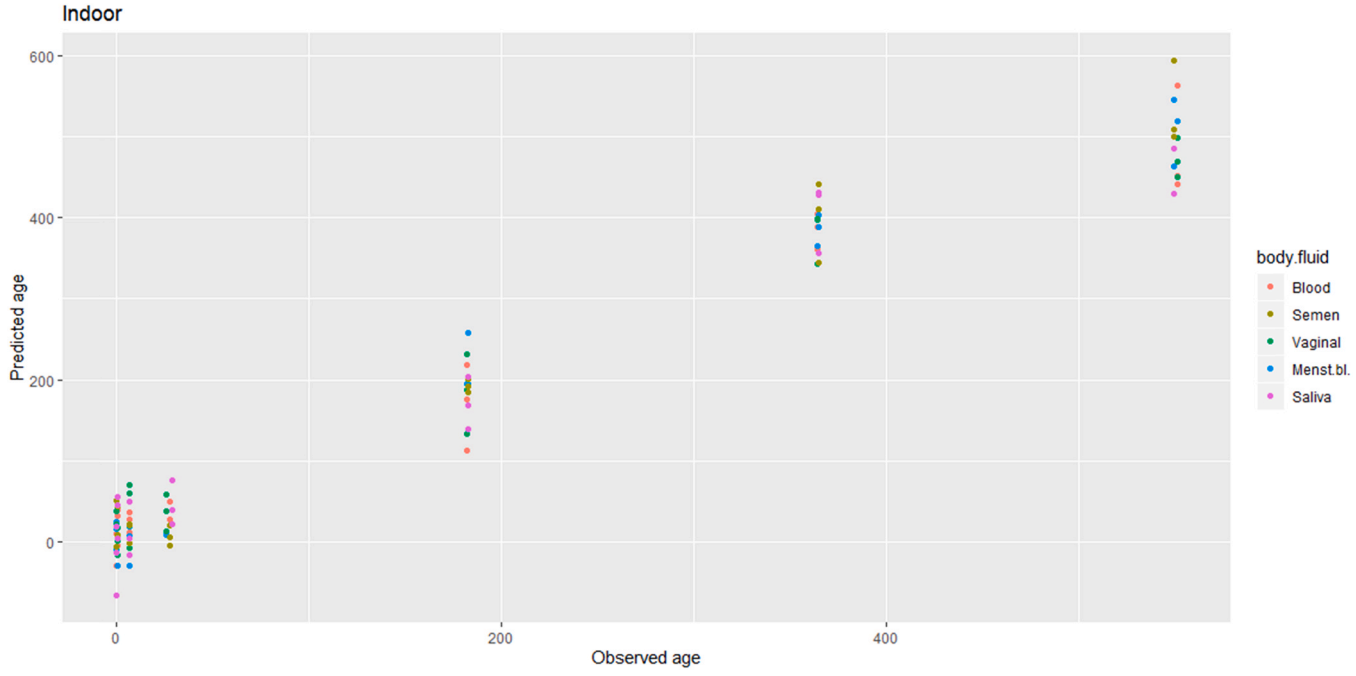

Fig. 7. Observed age of the indoor samples (y-axis) versus predicted (fitted) age (x-axis). The plot is based on the 32 taxa at order rank selected by lasso to be good age predictors. *(For interpretation of the references to color in this figure, the reader is referred to the online version of this article.)*

the park surrounding the university building, where the study was conducted. The presence of switch grass was unexpected at first because it is not native to the area, but we discovered that it is commonly utilized for decorative purposes in gardens. The body fluid samples were placed outdoors in March / April. Oak trees blossom around April to May, while switch grass starts to blossom in August. In addition, signatures of Olea (Olive tree) and Cercospora (ascomycete fungi) were identified. Interestingly, the pollen from Olea species has been found to travel long distances. For instance, *Olea europea* pollen most likely originating in Italy has been detected in Hungary [63]. Our results indicate, therefore, that outdoor samples contain not only environmental RNA from the immediate vicinity but possibly also from more distant locations, for example as a result of pollen transportation from other areas. Overall, we hypothesize the increase in environmental RNA to be useful as a first indicator for sample age, particularly if signatures specific to the sample location can be identified. It is, however, important to note that these signatures may change across seasons. Overall, further studies are needed in order to verify these findings and also to further assess the effect of different seasons and conditions on the samples.

When focusing on the prokaryotic fraction, we found three phyla that were of particular interest: Proteobacteria, Firmicutes and Actinobacteria. In outdoor samples, all three were detected in all body fluids across time but their abundance was body fluid specific and subject to change over time (Fig. 3 and Supplementary Fig. 5). This finding could potentially be used to develop qPCR or targeted MPS assays, where 16S-rRNA gene regions from these three phyla are analyzed. By computing the relative fractions of Proteobacteria, Firmicutes and Actinobacteria in the stain, its age might be estimated. From a forensic point of view, a targeted approach rather than a whole transcriptome approach would be beneficial, given that targeted sequencing requires less biological material, is cheaper, less labor intensive, and computationally simpler. Therefore, our observations will have to be confirmed with a targeted assay using more samples. Importantly, it will be necessary to determine whether the patterns hold across studies and geographical locations.

Compared to the outdoor samples, the composition of indoor samples showed less variation, both at the domain level (Fig. 5) as well as at the phylum level (Fig. 6), across the entire 1.5 year period. Our results are in agreement with a previous study that examined forensically relevant body fluid samples exposed to indoor conditions for 4 weeks. The study found that the exposed samples continue to harbor a characteristic microbial composition allowing for body fluid determination [16]. Our study suggests that the microbial composition is stable for a much longer period of time, thus providing greater evidence for the utility of bacterial signatures as an alternative approach for body fluid identification [3,64].

While bacterial signatures are overall stable, there are changes in the prokaryotic composition nonetheless. In this study we carried out Lasso regression analyses for indoor and outdoor samples to identify the bacterial orders changing most over time. A total of 26 taxa for outdoor samples (Fig. 4 and Supplementary Table 1) and 32 taxa for indoor samples (Fig. 7 and Supplementary Table 2) were found to be the smallest subsets of predictive taxa. Closer analysis of the selected taxa for both conditions revealed that they most probably originate from the environment, and increase in abundance over time in the samples. A large proportion of the taxa identified are characteristic of soil. Others were found to be from the Holosporales order, which comprises intracellular bacteria associated with protists and invertebrates. We also found small numbers of human, animal or plant pathogens (e.g. Legionellales or Burkholderiales). In addition to terrestrial bacteria, we also found taxa associated with marine environments. While this phenomenon may seem surprising, it is not unknown. For example, Evans et al. [65] were able to detect marine microbes in the fog of the Namib Desert (Namibia), which they explained was a result of the potential of fog to act as a transportation vector for microbial species. Given that several studies have demonstrated that bacteria can propagate by air over long distances [66,67], it is not unusual to detect their presence in the outdoor samples. In indoor samples too we find environmental bacteria. These may not necessarily be characteristic of a laboratory; rather, the bacteria may have entered the building from outside or have their origin in the soil of different plants kept in the room. It has been shown that airborne bacteria from the outdoor environment can be introduced into a house through open windows and doors as well as the ventilation system [68]. For example, in a study of indoor bacterial communities by Meadow et al., the authors detected changes that followed those occurring in outdoor communities, albeit after a time lag that was dependent on the ventilation (mechanical or natural) of the room [69]. In addition, we suspect that the detected changes in both outdoor and indoor samples are, to a certain extent, site-specific. Therefore, further experiments have to be performed to gain more insights into the time dependent changes of microbial communities at different locations. Interestingly, the small number of bacterial RNA reads found in the negative controls also suggests that bacteria transported through air deposit on samples. The indoor and outdoor negative controls contained environmental taxa that were also found on the aged body fluid samples in that specific condition.

When taking into account the different domains, we found a very prominent time effect in the outdoor samples, which was due to a shift from prokaryotes to non-human eukaryotes after 6 months. After this time point, the samples were dominated by the genera Quercus and Panicum, which are characteristic for the environment the samples were exposed to in this study. The composition of indoor samples showed less variation, both at the domain level as well as at the phylum level, across the entire 1.5 year period. Nevertheless, for both indoor and outdoor samples, we were able to identify bacterial orders suitable as predictors for TsD. We would like to highlight that the observed changes across time are not due to changes in bacteria specific to a certain body fluid but due to accumulation of bacteria from the environment. Therefore, the age of a stain may be determined by analyzing the accumulation of environmental RNA, which distorts the taxonomic composition of the body fluids. Given that we have good knowledge on taxa specific to human body fluids, we may in principle be able to detect this change, and use it as a "clock" to determine TsD. A similar approach is used to diagnose patients suffering from inflammatory bowel disease (IBD), where the deviation from a "normal" healthy gut microbiome (dysbiosis) is detected [70]. Such an approach could also be used to estimate the age of a stain.

Further studies may be able to reveal patterns in the changes of endogenous bacteria of body fluids across time. However, such patterns may be more complex, depending heavily on temperature, humidity and other conditions. Degradation in itself may be very informative for TsD, when it leads to consistent time-dependent patterns. In a recent analysis of the degradation pattern of human mRNA in forensically relevant body fluids, the researchers found that the 5'end of the transcripts degraded faster than the 3'end [21]. A qPCR approach assessing the degradation pattern of four selected blood transcripts showed promising results in determining TsD. It was shown that the age of a sample can be estimated with an accuracy of about two to four weeks, if the sample is younger than six months, while samples with an age of six month to one year can be estimated with an accuracy of four to six weeks. Hence, in depth analyses of the degradation pattern of bacterial RNA transcripts originating from body fluid specific bacteria might be a potential approach to look for time dependent changes and TsD estimation.

In this study, the Lasso regression analyses were conducted across all body fluids, enabling the selection of a set of bacterial orders useful for TsD estimation across different sample types. Thus, prediction could be conducted without prior knowledge of the bodily origin of the stain. It would be interesting, however, to conduct these regression analyses on each body fluid individually, for application to cases where the bodily origin of the stain is known. These analyses would require larger sample sizes and potentially also data from diverse research groups and geographical locations, in order to avoid overfitting.

5. Conclusion

The aim of our study was to determine the utility of microbial RNA markers for TsD estimation. We found that the observed changes in the microbial composition of body fluid stains across time are mainly due to accumulation of bacteria from the environment rather than to changes associated with body fluid specific bacteria. However, both aspects could be used for TsD estimation. Orders of these environmental bacteria could be evaluated for TsD prediction, considering the location and environment of the crime scene. In addition, the degradation pattern of bacterial RNA transcripts originating from body fluid specific bacteria might be a potential approach to look for time dependent changes and TsD estimation. The data generated within the present study will serve as a basis to further evaluate the effect of different environments and locations on the microbial composition of forensically relevant body fluids. Ultimately, we aim to establish simple methods (e.g. qPCR or targeted MPS) for the estimation of TsD in forensic stains. Based on our previous study [50], RNA-Seq data is also suitable for body fluid identification. Therefore, microbial RNA analysis would allow for simultaneous TsD prediction as well as body fluid identification. It has previously been shown that DNA and RNA can be co-extracted from the same samples [71–75]. In addition, this method can be applied to casework samples to simultaneously identify the donor and the origin of a stain [76]. Hence, we envision that once suitable markers have been selected, individual identification, body fluid identification and TsD determination should be possible out of the same stain material.

Acknowledgments

This project was supported by the Swiss National Science Foundation (SNF, project-Nr. 310030L_175854). Special thanks to Corinne Moser for technical assistance and to all volunteers who donated samples.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.fsigen.2021.102537.

References

[1] K. Virkler, I.K. Lednev, Analysis of body fluids for forensic purposes: from laboratory testing to non-destructive rapid confirmatory identification at a crime scene, Forensic Sci. Int. 188 (1–3) (2009) 1–17, https://doi.org/10.1016/j.forsciint.2009.02.013.
[2] J.H. An, K.J. Shin, W.I. Yang, H.Y. Lee, Body fluid identification in forensics, BMB Rep. 45 (10) (2012) 545–553.
[3] C. Haas, E. Hanson, J. Ballantyne, mRNA and MicroRNA for Body Fluid Identification, in: J. Siegel, P. Saukko (Eds.), Encyclopedia of Forensic Sciences, second ed., Academic Press, Waltham, 2013, pp. 402–408.
[4] J. Juusola, J. Ballantyne, Messenger RNA profiling: a prototype method to supplant conventional methods for body fluid identification, Forensic Sci. Int. 135 (2) (2003) 85–96.
[5] J. Juusola, J. Ballantyne, Multiplex mRNA profiling for the identification of body fluids, Forensic Sci. Int. 152 (1) (2005) 1–12, https://doi.org/10.1016/j.forsciint.2005.02.020.
[6] C. Haas, B. Klesser, C. Maake, W. Bar, A. Kratzer, mRNA profiling for body fluid identification by reverse transcription endpoint PCR and realtime PCR, Forensic Sci. Int. Genet. 3 (2) (2009) 80–88, https://doi.org/10.1016/j.fsigen.2008.11.003.
[7] R.I. Fleming, S. Harbison, The development of a mRNA multiplex RT-PCR assay for the definitive identification of body fluids, Forensic Sci. Int. Genet. 4 (4) (2010) 244–256, https://doi.org/10.1016/j.fsigen.2009.10.006.
[8] J. Juusola, J. Ballantyne, mRNA profiling for body fluid identification by multiplex quantitative RT-PCR, J. Forensic Sci. 52 (6) (2007) 1252–1262, https://doi.org/10.1111/j.1556-4029.2007.00550.x.
[9] I. Gomes, F. Kohlmeier, P. Schneider, Genetic markers for body fluid and tissue identification in forensics, Forensic Sci. Int. Genet. Suppl. Ser. 3 (1) (2011) e469–e470.
[10] D. Frumkin, A. Wasserstrom, B. Budowle, A. Davidson, DNA methylation-based forensic tissue identification, Forensic Sci. Int. Genet. 5 (5) (2011) 517–524, https://doi.org/10.1016/j.fsigen.2010.12.001.
[11] A. Vidaki, F. Giangasparo, D. Syndercombe Court, Discovery of potential DNA methylation markers for forensic tissue identification using bisulphite pyrosequencing, Electrophoresis 37 (21) (2016) 2767–2779, https://doi.org/10.1002/elps.201600261.
[12] H.Y. Yang, B. Zhou, H.T. Deng, M. Prinz, D. Siegel, Body fluid identification by mass spectrometry, Int. J. Leg. Med. 127 (6) (2013) 1065–1077, https://doi.org/10.1007/s00414-013-0848-1.
[13] T.H. Clarke, A. Gomez, H. Singh, K.E. Nelson, L.M. Brinkac, Integrating the microbiome as a resource in the forensics toolkit, Forensic Sci. Int. Genet. 30 (2017) 141–147, https://doi.org/10.1016/j.fsigen.2017.06.008.
[14] E.N. Hanssen, E. Avershina, K. Rudi, P. Gill, L. Snipen, Body fluid prediction from microbial patterns for forensic application, Forensic Sci. Int. Genet. 30 (2017) 10–17, https://doi.org/10.1016/j.fsigen.2017.05.009.
[15] E.N. Hanssen, K.H. Liland, P. Gill, L. Snipen, Optimizing body fluid recognition from microbial taxonomic profiles, Forensic Sci. Int. Genet. 37 (2018) 13–20, https://doi.org/10.1016/j.fsigen.2018.07.012.
[16] A. Dobay, C. Haas, G. Fucile, N. Downey, H.G. Morrison, A. Kratzer, N. Arora, Microbiome-based body fluid identification of samples exposed to indoor conditions, Forensic Sci. Int. Genet. 40 (2019) 105–113, https://doi.org/10.1016/j.fsigen.2019.02.010.
[17] R.H. Bremmer, K.G. de Bruin, M.J. van Gemert, T.G. van Leeuwen, M.C. Aalders, Forensic quest for age determination of bloodstains, Forensic Sci. Int. 216 (1–3) (2012) 1–11, https://doi.org/10.1016/j.forsciint.2011.07.027.
[18] S. Anderson, B. Howard, G.R. Hobbs, C.P. Bishop, A method for determining the age of a bloodstain, Forensic Sci. Int. 148 (1) (2005) 37–45, https://doi.org/10.1016/j.forsciint.2004.04.071.
[19] M. Bauer, S. Polzin, D. Patzelt, Quantification of RNA degradation by semi-quantitative duplex and competitive RT-PCR: a possible indicator of the age of bloodstains? Forensic Sci. Int. 138 (1–3) (2003) 94–103.
[20] S.E. Anderson, G.R. Hobbs, C.P. Bishop, Multivariate analysis for estimating the age of a bloodstain, J. Forensic Sci. 56 (1) (2011) 186–193, https://doi.org/10.1111/j.1556-4029.2010.01551.x.
[21] J. Fu, R.W. Allen, A method to estimate the age of bloodstains using quantitative PCR, Forensic Sci. Int. Genetic. 39 (2019) 103–108, https://doi.org/10.1016/j.fsigen.2018.12.004.
[22] S. Alshehhi, P.R. Haddrill, Estimating time since deposition using quantification of RNA degradation in body fluid-specific markers, Forensic Sci. Int. 298 (2019) 58–63, https://doi.org/10.1016/j.forsciint.2019.02.046.
[23] M.H. Lin, D.F. Jones, R. Fleming, Transcriptomic analysis of degraded forensic body fluids, Forensic Sci. Int. Genet. 17 (2015) 35–42, https://doi.org/10.1016/j.fsigen.2015.03.005.
[24] K.D. Weinbrecht, J. Fu, M.E. Payton, R.W. Allen, Time-dependent loss of mRNA transcripts from forensic stains, Res. Rep. Forensic Med. Sci. 7 (2017) 1–12, https://doi.org/10.2147/RRFMS.S125782.
[25] M. Oliveira, A. Amorim, Microbial forensics: new breakthroughs and future prospects, Appl. Microbiol. Biotechnol. 102 (24) (2018) 10377–10391, https://doi.org/10.1007/s00253-018-9414-6.
[26] S. Harbison, R. Flemming, Forensic body fluid identification: state of the art, Res. Rep. Forensic Med. Sci. 6 (2016) 11–23, https://doi.org/10.2147/RRFMS.S57994.
[27] J.L. Metcalf, Z.Z. Xu, A. Bouslimani, P. Dorrestein, D.O. Carter, R. Knight, Microbiome tools for forensic science, Trends Biotechnol. 35 (9) (2017) 814–823, https://doi.org/10.1016/j.tibtech.2017.03.006.
[28] R. Sender, S. Fuchs, R. Milo, Revised estimates for the number of human and bacteria cells in the body, PLoS Biol. 14 (8) (2016), 1002533, https://doi.org/10.1371/journal.pbio.1002533.
[29] R. Sender, S. Fuchs, R. Milo, Are we really vastly outnumbered? Revisiting the ratio of bacterial to host cells in humans, Cell 164 (3) (2016) 337–340, https://doi.org/10.1016/j.cell.2016.01.013.
[30] T. Sijen, Molecular approaches for forensic cell type identification: on mRNA, miRNA, DNA methylation and microbial markers, Forensic Sci. Int. Genet. 18 (2015) 21–32, https://doi.org/10.1016/j.fsigen.2014.11.015.
[31] A.D. Roeder, C. Haas, mRNA profiling using a minimum of five mRNA markers per body fluid and a novel scoring method for body fluid identification, Int. J. Leg. Med. 127 (4) (2013) 707–721, https://doi.org/10.1007/s00414-012-0794-3.
[32] R.I. Fleming, S. Harbison, The use of bacteria for the identification of vaginal secretions, Forensic Sci. Int. Genet. 4 (5) (2010) 311–315, https://doi.org/10.1016/j.fsigen.2009.11.008.
[33] T. Akutsu, H. Motani, K. Watanabe, H. Iwase, K. Sakurada, Detection of bacterial 16S ribosomal RNA genes for forensic identification of vaginal fluid, Leg. Med. 14 (3) (2012) 160–162, https://doi.org/10.1016/j.legalmed.2012.01.005.
[34] D.A. Power, S.J. Cordiner, J.A. Kieser, G.R. Tompkins, J. Horswell, PCR-based detection of salivary bacteria as a marker of expirated blood, Sci. Justice 50 (2) (2010) 59–63, https://doi.org/10.1016/j.scijus.2009.04.006.
[35] H. Nakanishi, A. Kido, T. Ohmori, A. Takada, M. Hara, N. Adachi, K. Saito, A novel method for the identification of saliva by detecting oral streptococci using PCR, Forensic Sci. Int. 183 (1–3) (2009) 20–23, https://doi.org/10.1016/j.forsciint.2008.10.003.
[36] H. Nakanishi, H. Shojo, T. Ohmori, M. Hara, A. Takada, N. Adachi, K. Saito, Identification of feces by detection of Bacteroides genes, Forensic Sci. Int. Genet. 7 (1) (2013) 176–179, https://doi.org/10.1016/j.fsigen.2012.09.006.
[37] Human Microbiome Project Consortium, Structure, function and diversity of the healthy human microbiome, Nature 486 (7402) (2012) 207–214, https://doi.org/10.1038/nature11234.
[38] H.M.P. Integrative, Research Network Consortium, The Integrative Human Microbiome Project, Nature 569 (7758) (2019) 641–648, https://doi.org/10.1038/s41586-019-1238-8.
[39] H.M.P. Integrative, Research Network Consortium, The Integrative Human Microbiome Project: dynamic analysis of microbiome-host omics profiles during periods of human health and disease, Cell Host Microbe 16 (3) (2014) 276–289, https://doi.org/10.1016/j.chom.2014.08.014.

[40] J. Lloyd-Price, A. Mahurkar, G. Rahnavard, J. Crabtree, J. Orvis, A.B. Hall, A. Brady, H.H. Creasy, C. McCracken, M.G. Giglio, D. McDonald, E.A. Franzosa, R. Knight, O. White, C. Huttenhower, Erratum: Strains, functions and dynamics in the expanded Human Microbiome Project, Nature 551 (7679) (2017) 256, https://doi.org/10.1038/nature24485.
[41] I. Cho, M.J. Blaser, The human microbiome: at the interface of health and disease, Nat. Rev. Genet. 13 (4) (2012) 260–270, https://doi.org/10.1038/nrg3182.
[42] J.A. Gilbert, M.J. Blaser, J.G. Caporaso, J.K. Jansson, S.V. Lynch, R. Knight, Current understanding of the human microbiome, Nat. Med. 24 (4) (2018) 392–400, https://doi.org/10.1038/nm.4517.
[43] J. Tackmann, N. Arora, T.S.B. Schmidt, J.F.M. Rodrigues, C. von Mering, Ecologically informed microbial biomarkers and accurate classification of mixed and unmixed samples in an extensive cross-study of human body sites, Microbiome 6 (1) (2018) 192, https://doi.org/10.1186/s40168-018-0565-6.
[44] L.A. David, C.F. Maurice, R.N. Carmody, D.B. Gootenberg, J.E. Button, B.E. Wolfe, A.V. Ling, A.S. Devlin, Y. Varma, M.A. Fischbach, S.B. Biddinger, R.J. Dutton, P. J. Turnbaugh, Diet rapidly and reproducibly alters the human gut microbiome, Nature 505 (7484) (2014) 559–563, https://doi.org/10.1038/nature12820.
[45] E.K. Costello, C.L. Lauber, M. Hamady, N. Fierer, J.I. Gordon, R. Knight, Bacterial community variation in human body habitats across space and time, Science 326 (5960) (2009) 1694–1697, https://doi.org/10.1126/science.1177486.
[46] J.G. Caporaso, C.L. Lauber, E.K. Costello, D. Berg-Lyons, A. Gonzalez, J. Stombaugh, D. Knights, P. Gajer, J. Ravel, N. Fierer, J.I. Gordon, R. Knight, Moving pictures of the human microbiome, Genome Biol. 12 (5) (2011) 50, https://doi.org/10.1186/gb-2011-12-5-r50.
[47] C. Díez López, D. Montiel González, C. Haas, A. Vidaki, M. Kayser, Microbiome-based body site of origin classification of forensically relevant blood traces, Forensic Sci. Int. Genet. 47 (2020), 102280, https://doi.org/10.1016/j.fsigen.2020.102280.
[48] C. Díez López, A. Vidaki, A. Ralf, D. Montiel González, D. Radjabzadeh, R. Kraaij, A.G. Uitterlinden, C. Haas, O. Lao, M. Kayser, Novel taxonomy-independent deep learning microbiome approach allows for accurate classification of different forensically relevant human epithelial materials, Forensic Sci. Int. Genet. 41 (2019) 72–82, https://doi.org/10.1016/j.fsigen.2019.03.015.
[49] F.C.A. Quaak, Y. van de Wal, P.A. Maaskant-van Wijk, I. Kuiper, Combining human STR and microbial population profiling: Two case reports, Forensic Sci. Int.: Genet. 37 (2018) 196–199, https://doi.org/10.1016/j.fsigen.2018.08.018.
[50] A.P. Salzmann, G. Russo, S. Aluri, C. Haas, Transcription and microbial profiling of body fluids using a massively parallel sequencing approach, Forensic Sci. Int. Genet. 43 (2019), 102149, https://doi.org/10.1016/j.fsigen.2019.102149.
[51] A.P. Salzmann, G. Russo, S. Kreutzer, C. Haas, Degradation of human mRNA transcripts over time as an indicator of the time since deposition (TsD) in biological crime scenetraces, Forensic Sci. Int. Genet. 53 (2021), https://doi.org/10.1016/j.fsigen.2021.102524.
[52] S. Finnegan, S.L. Percival, EDTA: an antimicrobial and antibiofilm agent for use in wound care, Adv. Wound Care 4 (7) (2015) 415–421, https://doi.org/10.1089/wound.2014.0577.
[53] A. Umerska, M. Strandh, V. Cassisa, N. Matougui, M. Eveillard, P. Saulnier, Synergistic effect of combinations containing EDTA and the antimicrobial peptide AA230, an Arenicin-3 derivative, on gram-negative bacteria, Biomolecules 8 (4) (2018) 122, https://doi.org/10.3390/biom8040122.
[54] Tecan Genomics, Trio RNA-Seq - Product Sheet, 2019. https://www.nugen.com/sites/default/files/401141_v1_Product_Sheet%3A_Trio_RNA-Seq_4145.pdf (Accessed 14 October 2020).
[55] D.E. Wood, J. Lu, B. Langmead, Improved metagenomic analysis with Kraken 2, Genome Biol. 20 (1) (2019) 257, https://doi.org/10.1186/s13059-019-1891-0.
[56] G.B. Gloor, J.R. Wu, V. Pawlowsky-Glahn, J.J. Egozcue, It's all relative: analyzing microbiome data as compositions, Ann. Epidemiol. 26 (5) (2016) 322–329, https://doi.org/10.1016/j.annepidem.2016.03.003.
[57] T. Hastie, R. Tibshirani, R.J. Tibshirani, Extended Comparisons of Best Subset Selection, Forward Stepwise Selection, and the Lasso, 2017. https://arxiv.org/abs/1707.08692 (Accessed 30 September 2020).
[58] J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for generalized linear models via coordinate descent, J. Stat. Softw. 33 (1) (2010) 1–22.
[59] N. Simon, J. Friedman, T. Hastie, R. Tibshirani, Regularization paths for Cox's proportional hazards model via coordinate descent, J. Stat. Softw. 39 (5) (2011) 1–13, https://doi.org/10.18637/jss.v039.i05.
[60] R. Tibshirani, J. Bien, J. Friedman, T. Hastie, N. Simon, J. Taylor, R.J. Tibshirani, Strong rules for discarding predictors in lasso-type problems, J. R. Stat. Soc. Ser. B Stat. Method. 74 (2) (2012) 245–266, https://doi.org/10.1111/j.1467-9868.2011.01004.x.
[61] F. Kohlmeier, P.M. Schneider, Successful mRNA profiling of 23 years old blood stains, Forensic Sci. Int. Genet. 6 (2) (2012) 274–276, https://doi.org/10.1016/j.fsigen.2011.04.007.
[62] M. Hara, H. Nakanishi, S. Takahashi, A. Kido, K. Saito, A. Takada, STR and Y-STR genotyping of 30–50-year-old semen stains, Forensic Sci. Int. Genet. Suppl. Ser. 4 (1) (2013) e99–e100, https://doi.org/10.1016/j.fsigss.2013.10.051.
[63] E. Tedeschini, O. Udvardy, M. Sofiev, J. Palamarchuk, L. Makra, D. Magyar, Long distance transport of Olea europea pollen over Central Europe, in: C. De Linares, J. Belmonte (Eds.), Mediterranean Palynology Symposium-Abstracts Book, La Imprenta Comunicación Gráfica s.l, Barcelona, Spain, 2017, p. 134.
[64] A. Lindenbergh, P. Maaskant, T. Sijen, Implementation of RNA profiling in forensic casework, Forensic Sci. Int. Genet. 7 (1) (2013) 159–166, https://doi.org/10.1016/j.fsigen.2012.09.003.
[65] S.E. Evans, M.E. Dueker, J.R. Logan, K.C. Weathers, The biology of fog: results from coastal Maine and Namib Desert reveal common drivers of fog microbial composition, Sci. Total Environ. 647 (2019) 1547–1556, https://doi.org/10.1016/j.scitotenv.2018.08.045.
[66] N. Els, C. Larose, K. Baumann-Stanzer, R. Tignat-Perrier, C. Keuschnig, T.M. Vogel, B. Sattler, Microbial composition in seasonal time series of free tropospheric air and precipitation reveals community separation, Aerobiologia 35 (4) (2019) 671–701, https://doi.org/10.1007/s10453-019-09606-x.
[67] E. Innocente, S. Squizzato, F. Visin, C. Facca, G. Rampazzo, V. Bertolini, I. Gandolfi, A. Franzetti, R. Ambrosini, G. Bestetti, Influence of seasonality, air mass origin and particulate matter chemical composition on airborne bacterial community structure in the Po Valley, Italy, Sci. Total Environ. 593–594 (2017) 677–687, https://doi.org/10.1016/j.scitotenv.2017.03.199.
[68] S. Fujiyoshi, D. Tanaka, F. Maruyama, Transmission of airborne bacteria across built environments and its measurement standards: a review, Front. Microbiol. 8 (2336) (2017) 2336, https://doi.org/10.3389/fmicb.2017.02336.
[69] J.F. Meadow, A.E. Altrichter, S.W. Kembel, J. Kline, G. Mhuireach, M. Moriyama, D. Northcutt, T.K. O'Connor, A.M. Womack, G.Z. Brown, J.L. Green, B.J. M. Bohannan, Indoor airborne bacterial communities are influenced by ventilation, occupancy, and outdoor air source, Indoor Air 24 (1) (2014) 41–48, https://doi.org/10.1111/ina.12047.
[70] C. Casén, H.C. Vebø, M. Sekelja, F.T. Hegge, M.K. Karlsson, E. Ciemniejewska, S. Dzankovic, C. Frøyland, R. Nestestog, L. Engstrand, P. Munkholm, O.H. Nielsen, G. Rogler, M. Simrén, L. Öhman, M.H. Vatn, K. Rudi, Deviations in human gut microbiota: a novel diagnostic test for determining dysbiosis in patients with IBS or IBD, Aliment. Pharmacol. Ther. 42 (1) (2015) 71–83, https://doi.org/10.1111/apt.13236.
[71] M. Bauer, D. Patzelt, A method for simultaneous RNA and DNA isolation from dried blood and semen stains, Forensic Sci. Int. 136 (1–3) (2003) 76–78.
[72] C. Haas, E. Hanson, M.J. Anjos, W. Bar, R. Banemann, A. Berti, E. Borges, C. Bouakaze, A. Carracedo, M. Carvalho, V. Castella, A. Choma, G. De Cock, M. Dotsch, P. Hoff-Olsen, P. Johansen, F. Kohlmeier, P.A. Lindenbergh, B. Ludes, O. Maronas, D. Moore, M.L. Morerod, N. Morling, H. Niederstatter, F. Noel, W. Parson, G. Patel, C. Popielarz, E. Salata, P.M. Schneider, T. Sijen, B. Sviezena, M. Turanska, L. Zatkalikova, J. Ballantyne, RNA/DNA co-analysis from blood stains–results of a second collaborative EDNAP exercise, Forensic Sci. Int. Genet. 6 (1) (2012) 70–80, https://doi.org/10.1016/j.fsigen.2011.02.004.
[73] C. Haas, E. Hanson, M.J. Anjos, R. Banemann, A. Berti, E. Borges, A. Carracedo, M. Carvalho, C. Courts, G. De Cock, M. Dotsch, S. Flynn, I. Gomes, C. Hollard, B. Hjort, P. Hoff-Olsen, K. Hribikova, A. Lindenbergh, B. Ludes, O. Maronas, N. McCallum, D. Moore, N. Morling, H. Niederstatter, F. Noel, W. Parson, C. Popielarz, C. Rapone, A.D. Roeder, Y. Ruiz, E. Sauer, P.M. Schneider, T. Sijen, D. S. Court, B. Sviezena, M. Turanska, A. Vidaki, L. Zatkalikova, J. Ballantyne, RNA/DNA co-analysis from human saliva and semen stains–results of a third collaborative EDNAP exercise, Forensic Sci. Int. Genet. 7 (2) (2013) 230–239, https://doi.org/10.1016/j.fsigen.2012.10.011.
[74] C. Haas, E. Hanson, M.J. Anjos, K.N. Ballantyne, R. Banemann, B. Bhoelai, E. Borges, M. Carvalho, C. Courts, G. De Cock, K. Drobnic, M. Dotsch, R. Fleming, C. Franchi, I. Gomes, G. Hadzic, S.A. Harbison, J. Harteveld, B. Hjort, C. Hollard, P. Hoff-Olsen, C. Huls, C. Keyser, O. Maronas, N. McCallum, D. Moore, N. Morling, H. Niederstatter, F. Noel, W. Parson, C. Phillips, C. Popielarz, A.D. Roeder, L. Salvaderi, E. Sauer, P.M. Schneider, G. Shanthan, D.S. Court, M. Turanska, R. A. van Oorschot, M. Vennemann, A. Vidaki, L. Zatkalikova, J. Ballantyne, RNA/DNA co-analysis from human menstrual blood and vaginal secretion stains: results of a fourth and fifth collaborative EDNAP exercise, Forensic Sci. Int. Genet. 8 (1) (2014) 203–212, https://doi.org/10.1016/j.fsigen.2013.09.009.
[75] C. Haas, E. Hanson, W. Bar, R. Banemann, A.M. Bento, A. Berti, E. Borges, C. Bouakaze, A. Carracedo, M. Carvalho, A. Choma, M. Dotsch, M. Duriancikova, P. Hoff-Olsen, C. Hohoff, P. Johansen, P.A. Lindenbergh, B. Loddenkotter, B. Ludes, O. Maronas, N. Morling, H. Niederstatter, W. Parson, G. Patel, C. Popielarz, E. Salata, P.M. Schneider, T. Sijen, B. Sviezena, L. Zatkalikova, J. Ballantyne, mRNA profiling for the identification of blood–results of a collaborative EDNAP exercise, Forensic Sci. Int. Genet. 5 (1) (2011) 21–26, https://doi.org/10.1016/j.fsigen.2010.01.003.
[76] A.P. Salzmann, M. Bamberg, C. Courts, G. Dorum, A. Gosch, T. Hadrys, G. Hadzic, M. Neis, P.M. Schneider, T. Sijen, M.V. den Berge, P. Wiegand, C. Haas, mRNA profiling of mock casework samples: results of a FoRNAP collaborative exercise, Forensic Sci. Int. Genet. 50 (2021), 102409, https://doi.org/10.1016/j.fsigen.2020.102409.